

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

THE NEW YORK TIMES COMPANY,

Plaintiff,

v.

MICROSOFT CORPORATION, OPENAI, INC.,
OPENAI LP, OPENAI GP, LLC, OPENAI, LLC,
OPENAI OPCO LLC, OPENAI GLOBAL LLC,
OAI CORPORATION, LLC, and OPENAI
HOLDINGS, LLC,

Defendants.

Civil Action No. 1:23-cv-11195-SHS

**PLAINTIFF'S RESPONSE TO DEFENDANTS' OBJECTIONS TO NONDISPOSITIVE
ORDER PURSUANT TO FEDERAL RULE OF CIVIL PROCEDURE 72(a)
(DKTS. 368, 371)**

TABLE OF CONTENTS

| | |
|--|----|
| INTRODUCTION | 1 |
| BACKGROUND | 3 |
| I. The Times Already Agreed to Provide Detailed Information About Its Revenue, Subscriptions, Web Traffic, and Licensing Efforts. | 3 |
| II. Judge Wang’s Discovery Orders Are Consistent with Governing Law. | 5 |
| A. Judge Wang’s November 22, 2024 Order. | 6 |
| B. Judge Wang’s December 2, 2024 Orders. | 7 |
| LEGAL STANDARD | 8 |
| ARGUMENT | 8 |
| I. The Court Correctly Denied Defendants’ Irrelevant and Overbroad Motions for All Documents Regarding Times Traffic, Subscribers, and Revenue..... | 9 |
| A. Defendants’ Requests Are Neither Relevant Nor Proportional to Their Fair Use Defense. | 9 |
| B. Microsoft’s Request for <i>All</i> Surveys Regarding Subscription Cancellations Is Overbroad and Irrelevant. | 12 |
| C. <i>Google</i> Does Not Support Compelling Production of These Documents. | 13 |
| II. The Court Correctly Denied Defendants’ Licensing Requests, Which Are Overbroad and Irrelevant to the Market Substitution Inquiry | 14 |
| III. Microsoft’s Remaining Arguments Are Unavailing. | 16 |
| A. Microsoft’s Requests are Not Relevant to Assessing the Transformativeness of Microsoft’s Unauthorized Use. | 16 |
| B. Microsoft’s Requests are Not Relevant to Show Substantial Non-Infringing Use. | 17 |
| CONCLUSION..... | 19 |

TABLE OF AUTHORITIES

| | Page(s) |
|---|----------------|
| Cases | |
| <i>Am. Geophysical Union v. Texaco Inc.</i> , 60 F.3d 913 (2d Cir. 1994)..... | 10 |
| <i>Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith</i> , 598 U.S. 508 (2023)..... | 6 |
| <i>Arista Recs. LLC v. Usenet.com, Inc.</i> , 633 F. Supp. 2d 124 (S.D.N.Y. 2009)..... | 18 |
| <i>Authors Guild v. Google, Inc.</i> , 804 F.3d 202 (2d Cir. 2015)..... | 10 |
| <i>Coventry Cap. US LLC v. EEA Life Settlements Inc.</i> , 333 F.R.D. 60 (S.D.N.Y. 2019) | 8, 13 |
| <i>Diabat v. Credit Suisse Grp. AG</i> , 2024 4252502 (S.D.N.Y. Sept. 19, 2024)..... | 17 |
| <i>EMI Christian Music Grp., Inc. v. MP3Tunes, LLC</i> , 844 F.3d 79 (2d Cir. 2016)..... | 17 |
| <i>Fashion Exch. LLC v. Hybrid Promotions, LLC</i> , No. 14-CV-1254 (SHS), 2021 WL 1172265 (S.D.N.Y. Mar. 29, 2021)..... | 2 |
| <i>Google LLC v. Oracle Am., Inc.</i> , 593 U.S. 1 (2021)..... <i>passim</i> | |
| <i>Hachette Book Grp., Inc. v. Internet Archive</i> , 115 F.4th 163 (2d Cir. 2024) | 10, 11, 15 |
| <i>Kumaran v. Vision Fin. Markets, LLC</i> , 2022 WL 17540669 (S.D.N.Y. Dec. 6, 2022) | 8 |
| <i>Madi v. United Parcel Serv. Gen. Serv. Co.</i> , 1999 WL 193403 (S.D.N.Y. Apr. 7, 1999)..... | 8 |
| <i>Nikkal Indus., Ltd. v. Salton, Inc.</i> , 689 F. Supp. 187 (S.D.N.Y. 1988) | 8 |
| <i>R.F.M.A.S., Inc. v. So</i> , 748 F. Supp. 2d 244 (S.D.N.Y. 2010)..... | 2, 8 |

| | |
|--|--------|
| <i>Sony Corp. of America v. Universal City Studios, Inc.,</i> 464 U.S. 417 (1984)..... | 17, 18 |
| <i>U2 Home Ent., Inc. v. Hong Wei Int'l Trading Inc.,</i> 2007 WL 2327068 (S.D.N.Y. Aug. 13, 2007)..... | 8 |
| <i>Viacom Int'l Inc. v. Youtube Inc.,</i> 253 F.R.D. 256 (S.D.N.Y. 2008) | 18 |
| <i>Winfield v. City of New York,</i> 2018 WL 2293070 (S.D.N.Y. May 18, 2018) | 11 |
| <i>Winfield v. City of New York,</i> 2018 WL 840085 (S.D.N.Y. Feb. 12, 2018)..... | 15, 16 |
| Statutes | |
| 17 U.S.C. § 107(4)..... | 10 |
| Rules | |
| Fed. R. Civ. P. 34..... | 11 |
| Fed. R. Civ. P. 72(a) | 8, 16 |

INTRODUCTION

Defendants seek to set aside five discovery orders issued by Judge Wang, all of which rejected Defendants' far-reaching discovery requests into categories of documents that are irrelevant and/or exceed the needs of this case. *See* Dkt. 369 (appealing Dkts. 351, 354, and 355); Dkt. 373 (appealing Dkts. 352 & 353). Contrary to Defendants' telling, and consistent with Judge Wang's findings, The New York Times Company ("The Times") has already agreed to produce documents in all categories at issue in Defendants' motions. The issue here is that, as has been the pattern in this case, Defendants are not satisfied with documents sufficient to put on their anticipated fair use defense. Rather, they think being sued entitles them to discovery into nearly every document from entire departments within The Times. As Judge Wang has now repeatedly had to rule, such broad requests are "neither relevant nor proportional to the needs of the case." Dkt. 344 at 2. By their Objection, Defendants fail to show that Judge Wang's carefully considered rulings were clearly erroneous or contrary to law.

Defendants take issue with The Times for focusing its search of responsive documents and data showing changes in traffic, subscriptions, and revenue caused by **Defendants'** generative AI tools. However, as Judge Wang already explained in her November 22, 2024, Order, the fair use factors require "scrutiny of a *defendant's* purported use of the copyrighted work." Dkt. 344 at 2. Accordingly, Judge Wang correctly rejected Defendants' efforts to compel production of documents that are neither relevant nor proportional to Defendants' fair use defense. *See* Dkts. 351-55. This ruling is bolstered by the substantial amount of discovery The Times has already agreed to produce. Indeed, contrary to Defendants' misrepresentation, *see* Dkt. 373 at 6, The Times **has agreed** to produce nonprivileged analyses and reports regarding whether decreases in traffic to Times websites were attributable to Defendants' generative AI products **or other market**

forces—the very evidence they claim to seek. *See* Dkt. 310-1 at 34. The Times has also agreed to produce, *inter alia*, (1) licensing agreements related to the asserted works, regardless of whether those licenses relate to generative AI; (2) documents sufficient to show all attempts to license the asserted works for training or use in generative AI; and (3) *all* communications related to executed and contemplated agreements to license Times works to train generative AI models.¹ *See* Dkt. 310-1 at 45–46. Defendants have failed to provide a satisfactory justification for why The Times’s extensive production is inadequate to assess harm to the market for The Times’s works.

A magistrate judge’s “determinations on discovery matters are entitled to substantial deference” and can only be disturbed by the district court if they are “clearly erroneous” or “contrary to the law.” *R.F.M.A.S., Inc. v. So*, 748 F. Supp. 2d 244, 248 (S.D.N.Y. 2010) (citation omitted). As this Court has explained, a decision is “clearly erroneous” where the reviewer “is left with the definite and firm conviction that a mistake has been committed.” *Fashion Exch. LLC v. Hybrid Promotions, LLC*, No. 14-CV-1254 (SHS), 2021 WL 1172265, at *1 (S.D.N.Y. Mar. 29, 2021) (citation omitted). The imposition of reasonable guardrails to prevent excessively broad discovery cannot reasonably create any such conviction. Similarly, a ruling is “contrary to law” when it “fails to apply or misapplies relevant statutes, case law or rules of procedure.” *Id.* (citation omitted). Defendants have cited no statute, case, or rule of procedure that entitles them to boundless discovery into areas for which The Times is already producing substantial evidence or that are far afield from any claim or defense in this case. Simply put, Defendants have not met the high threshold for reversal of Judge Wang’s orders. The Times respectfully requests the Court deny Defendants’ motions to do so.

¹ The Times has attached an appendix which more fully outlines the requests relevant to Defendants’ motions to compel and the responsive discovery The Times has agreed to produce. *See* Ex. 1.

BACKGROUND

I. The Times Already Agreed to Provide Detailed Information About Its Revenue, Subscriptions, Web Traffic, and Licensing Efforts.

OpenAI served requests seeking two broad categories of documents about The Times's market harm: (1) decreases in revenue and web traffic to The Times's platforms ("Revenue and Traffic Requests"), *see* Dkt. 310-1 at 31–361 and (2) effects on The Times's licensing efforts, including regarding the existence of a market for LLM training ("Licensing Requests"), *see id.* at 45–46. Microsoft went further with its two "catch-all" RFPs: "[d]ocuments concerning the market for the Copyrighted Works-In-Suit" and "[d]ocuments concerning the market for derivatives of Your Copyrighted Works-In-Suit." Dkt. 315-1 at 8–9.

In response to these requests, The Times long ago agreed to provide detailed information about just these categories, including its revenue, subscriptions, and web traffic, as well as documents concerning the impact of Defendants' products on the market for The Times's works. This latter agreement included documents showing the impact of Defendants' products on the market for the Times's works as compared to other market factors. Specifically, The Times agreed to produce non-privileged:

1. documents regarding any harm or injury alleged in this case, which includes documents relating to reductions in readership, revenue, or web traffic;
2. documents concerning the impact of Defendants' products on The Times's business and/or revenue;
3. documents concerning how Defendants have diverted Times readers and online traffic;
4. analyses and reports regarding factors giving rise to changes in traffic to Times websites, *including whether such changes are attributable to Defendants' generative AI products or other market forces*;

5. documents concerning any revenue The Times receives for the reproduction, distribution, display, or advertising of its copyrighted works;
6. documents sufficient to show The Times's monthly revenue, including from subscriptions, advertising, and affiliate links from 2015 to present;
7. documents regarding decreases in online traffic or impacts on Wirecutter subscriptions and advertising revenue attributed to Defendants;
8. documents sufficient to show total page views for the Asserted Works;
9. documents sufficient to show monthly user visits and traffic to The Times's websites from 2018 to present; and
10. documents sufficient to show the number of monthly Times subscribers from 2015 to present.

See Ex. 1.

When the above did not satisfy Defendants, The Times further agreed to produce (1) documents sufficient to show trends in readership for the Asserted Works and The Times's online subscriptions; and (2) documents sufficient to show trends in revenue generated in connection with content published on The Times's website. *See id.*, § A.

Defendants nevertheless pressed for yet more discovery. At bottom, the delta between the parties' positions was that, while The Times freely agreed to produce documents more than sufficient to show both the impact and source of the market harm, Defendants wanted every single document that could possibly speak to this issue.

On October 22, 2024, and November 18, 2024, Defendants sought pre-motion conferences before Judge Wang to compel further production, incorrectly claiming that The Times "refused to produce documents" regarding "the alleged decreases in traffic and revenue," Dkt. 310 at 2, when in fact The Times **had** agreed to produce analyses regarding changes in its audience and revenue. *See* Dkt. 276, 310, 315, 321. Again, the delta was Defendants' request for the kitchen sink. The

Times properly objected to Defendants' demand for all communications on these topics—a demand that would sweep in virtually every document generated by entire departments. As The Times explained to Defendants, these documents would have no bearing on Defendants' ability to assert a fair use defense and would be wholly disproportionate to the needs of this case.

Microsoft additionally sought far-reaching discovery into The Times's own use of Generative AI, which Judge Wang correctly ruled to be beyond the scope of the fair use inquiry. *See* Dkt. 344. Specifically, Microsoft requested that "the Court compel The Times to produce: (1) documents reflecting efforts by The Times to train, fine-tune, or otherwise develop and use a Generative AI tool; and (2) a survey conducted by The Times concerning workplace generative AI usage." Dkt. 321.

II. Judge Wang's Discovery Orders Are Consistent with Governing Law.

Judge Wang has devoted a substantial amount of time and attention to addressing discovery disputes in this case, which has been consolidated with two other cases brought by news publishers (the "News Cases"). By the time Judge Wang took up the instant disputes, Defendants had already litigated nine prior motions to compel before her, and the Court had already held four discovery hearings that consumed dozens of hours of Judge Wang's time. After careful consideration of the parties' arguments, Judge Wang denied each of Defendants' affirmative motions, including OpenAI's request that The Times produce every single reporters file it has ever generated—a request that, among numerous other issues, seeks documents wholly unrelated to the claims in this case. *See* Dkts. 243, 304, 365. Of particular relevance here, on November 22, 2024, Judge Wang issued a well-reasoned Opinion & Order rejecting OpenAI's motion to compel documents about the use of non-Defendant generative AI technology (the "November Order"). Dkt. 344. The

reasoning in that Order governed Judge Wang’s denial of Defendants’ instant motions to compel, which again are both overbroad and directed at irrelevant documents.

A. Judge Wang’s November 22, 2024 Order.

In the November Order, Judge Wang outlined the four statutory fair use factors and correctly noted that each factor “requires scrutiny of a *defendant’s* purported use of the copyrighted work(s), and whether that *defendant’s* use may constitute ‘fair use’ under the Act.” *Id.* at 2; *see also Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 527–28 (2023) (noting that in assessing the first statutory fair use factor, a court considers “**the copier’s** use of an original work” (emphasis added)); *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 35 (2021) (noting that in assessing the fourth statutory fair use factor, a court “must take into account the public benefits **the copying** will likely produce” (emphasis added)). In other words, the Court reasoned that it was not required to assess the benefits of generative AI generally, the effects of third-party generative AI tools on The Times’s business, or whether any Times’s employees use third-party generative AI tools at work. *See* Dkt. 344 at 5. The Court noted that its conclusion was bolstered by the fact that The Times’s Amended Complaint was “tightly focused on Defendant’s particular Gen AI products and [its] alleged use of The Times’s copyrighted content,” rather than alleging a “wholesale indictment of Gen AI tools,” as Defendant repeatedly claimed. *Id.* at 3.

The Court also rejected OpenAI’s reading of *Google*. *See id.* at 4. OpenAI contended (as Defendants do again here) that *Google* supported its position that “[e]vidence regarding [T]he Times’s creation, use, and positions on generative AI generally is directly relevant” to the first and fourth statutory fair use factors. *See* Dkt. 236 at 1. However, the Court correctly noted that “the Supreme Court suggested a more nuanced view of the market effects” inquiry; specifically, one that requires a court to assess the effects “**the copying** will likely produce,” Dkt. 344 at 4 (emphasis

added). Accordingly, the Court concluded that *Google* did “not support a finding of relevance” for the “broad scope of document production sought” by Defendants. *Id.* at 5.

B. Judge Wang’s December 2, 2024 Orders.

When Judge Wang issued the November Order, Defendants had several additional discovery motions pending, many of which overlapped with its prior—now denied—motions. *See* Dkt. 276 (OpenAI’s request for documents concerning (i) decreases in The Times’s subscriptions and web traffic, (ii) the impacts of defendants’ products on journalism, and (iii) The Times’s use of non-defendant products); Dkt. 310 (OpenAI’s request for documents and communications concerning (i) The Times’s market harm, including decreases in revenue and traffic; and (ii) the market for licensing Times works); Dkt. 315 (Microsoft’s request for, *inter alia*², documents concerning losses to Times (i) subscriptions, (ii) advertising revenue, and (iii) web traffic); Dkt. 320 (Microsoft’s request for documents concerning The Times’s use of non-defendant generative AI products); Dkt. 321 (same).

On December 2, 2024, the Court denied Defendants’ instant motions to compel, *see* Dkts. 351–355, “for the reasons stated” in the Court’s November Order. *See, e.g.*, Dkt. 351 (“ORDER denying [315] Letter Motion to Compel for the reasons stated in ECF 344.”). Although the November Order ruled on a separate motion to compel, the Court correctly held that the logic underlying that order was directly applicable to Defendants’ instant requests. Specifically, consistent with its November ruling, the Court again rejected Defendants’ expansive demand for every last document (in this case regarding market harm to The Times)—regardless of the reason for the harm—as it was “neither relevant nor proportional to the needs of the case.” *Id.*

² Dkt. 315 also sought to compel production of certain “high level economic and financial reporting data,” but the parties’ dispute on this request was resolved via agreement.

LEGAL STANDARD

Under Federal Rule of Civil Procedure 72(a), “a magistrate judge’s discovery ruling is a non-dispositive order which can only be overturned if it is ‘clearly erroneous or contrary to law.’” *Madi v. United Parcel Serv. Gen. Serv. Co.*, 1999 WL 193403, at *1 (S.D.N.Y. Apr. 7, 1999) (quoting Fed. R. Civ. P. 72(a)). “A finding is clearly erroneous if the reviewing court is left with the definite and firm conviction that a mistake has been committed.” *Coventry Cap. US LLC v. EEA Life Settlements Inc.*, 333 F.R.D. 60, 64 (S.D.N.Y. 2019) (quoting *R.F.M.A.S., Inc. v. So*, 748 F. Supp. 2d 244, 248 (S.D.N.Y. 2010)). “Similarly, a finding is contrary to law when it fails to apply or misapplies relevant statutes, case law, or rules of procedure.” *Id.* (quoting *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248).

“Therefore, pursuant to these standards, a magistrate judge’s determinations on discovery matters are entitled to substantial deference.” *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248; *see Nikkal Indus., Ltd. v. Salton, Inc.*, 689 F. Supp. 187, 189 (S.D.N.Y. 1988) (“Consistently, it has been held that a magistrate’s report resolving a discovery discourse between litigants should be afforded substantial deference and be overturned only if found to be an abuse of discretion.”). “The magistrate judge’s findings should not be rejected merely because the court would have decided the matter differently.” *Kumaran v. Vision Fin. Markets, LLC*, 2022 WL 17540669, at *6 (S.D.N.Y. Dec. 6, 2022). “The party seeking to overturn a magistrate judge’s decision thus carries a heavy burden.” *R.F.M.A.S., Inc.*, 748 F. Supp. 2d at 248 (quoting *U2 Home Ent., Inc. v. Hong Wei Int’l Trading Inc.*, 2007 WL 2327068, at *1 (S.D.N.Y. Aug. 13, 2007)).

ARGUMENT

Defendants have failed to satisfy their heavy burden of demonstrating that Judge Wang’s December 2, 2024, Orders were clearly erroneous or contrary to law. First, the Court correctly

denied Defendants' requests for all documents and communications concerning all causes of The Times's alleged decline in traffic, subscribers, and revenue, which are overbroad and irrelevant to the market substitution inquiry—that is, the fourth statutory fair use factor—particularly given the responsive discovery The Times has already agreed to provide. Second, the Court properly denied Defendants' requests for all internal communications or negotiations regarding licensing of the Asserted Works for purposes other than generative AI model training, which are similarly overbroad and irrelevant to the fourth statutory fair use factor. Third, the Court properly determined that The Times's use of, and positions about, “generative AI generally” were not relevant to assessing whether Microsoft's unauthorized use of The Times's works was transformative nor to Microsoft's purported non-infringing use defense, as reflected in her November Order. *See* ECF No. 344 at 5. As explained more fully below, Judge Wang's December 2, 2024, Orders should be affirmed.

I. The Court Correctly Denied Defendants' Irrelevant and Overbroad Motions for All Documents Regarding Times Traffic, Subscribers, and Revenue.

A. Defendants' Requests Are Neither Relevant Nor Proportional to Their Fair Use Defense.

First, Defendants seek all “documents and communications related to decreases in Times’ revenue[, traffic, and subscriptions], including the reasons for such decreases.” Dkt. 373 at 6. As Exhibit 1 makes clear, The Times provided reams of evidence responsive to this request. *See* Ex. 1, § A. For example, The Times agreed to provide analyses and reports regarding whether decreases in traffic to Times websites, if any, are attributable to GPT services *versus other market forces and factors*—*viz.*, the very evidence Defendants continue to insist the Times refuses to produce. *See* Dkt. 373 at 6. However, Defendants continue to claim The Times’s production is insufficient. At bottom, it appears Defendants will only be satisfied if the Times agrees to produce

every last document or email regarding changes in traffic, revenue, and subscriptions, regardless of the reasons for the fluctuation.

Defendants attempt to justify their requests by claiming that those documents and communications are relevant to the “substitution inquiry” of the fourth fair use factor. For the reasons already explained in Judge Wang’s November Order, this misunderstands the analysis. *See* Dkt. 344 (noting that the fair use factors “requires scrutiny of a **defendant’s** purported use of the copyrighted work(s), and whether that **defendant’s** use may constitute ‘fair use’ under the Act”); *cf. id.* at 5 (denying the broad scope of OpenAI’s requested production because it was “simply not relevant to [its] purported fair use defense”). The fourth fair use factor assesses “the effect of **the use** upon the potential market for or value of the copyrighted work.” 17 U.S.C. § 107(4) (emphasis added). In evaluating the fourth factor, courts determine “whether there was harm to the actual or potential markets for the copyrighted work.” *Google*, 593 U.S. at 26. More specifically, the fourth fair use factor focuses on “whether **the copy** brings to the marketplace a competing **substitute for the original, or its derivative**, so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may opt to acquire the copy in preference to the original.” *Authors Guild v. Google, Inc.*, 804 F.3d 202, 223 (2d Cir. 2015) (emphases added); *Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 189 (2d Cir. 2024)

Determining whether a use provides a “competing substitute” with a “likelihood” of substitution requires looking at the **infringing use itself**. *Id.*; *see, e.g., id.* at 224 (considering, when evaluating the fourth fair use factor, whether Google’s “snippet view” provided “significant portions of the original” to create a substitute); *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 927 (2d Cir. 1994) (noting that in evaluating a defendant’s fair use claim, “the analysis under the fourth factor must focus on the effect of [the defendant’s copying] upon the potential market

for or value of [the plaintiff's] individual articles"). In other words, assessing whether Defendants' use created a substitute requires looking, first and foremost, at the extent to which *their use* created verbatim copies that can serve as substitutes. As Judge Wang's November Order recognizes, that evidence will come primarily from Defendants, not The Times. *See Hachette Book Grp., Inc.*, 115 F.4th at 189 ("[T]he ultimate burden of proving that the secondary use does not compete in the relevant market is appropriately borne by the party asserting the [fair use] defense: the secondary user.").

Nevertheless, and as detailed above, The Times has already agreed to produce numerous categories of documents related to revenue, traffic, readership, and subscriptions, including analyses and reports regarding whether decreases in traffic to Times websites were attributable to Defendants' generative AI products *or other market forces*. *See* Ex. 1, § A. However, Defendants' expansive demand for every last document or email regarding changes in traffic to The Times's website or its revenue for any reason goes far beyond anything even possibly necessary or relevant to testing whether Defendants' infringing outputs could substitute for The Times's content. Defendants' request would capture a vast amount of trivial daily communications between Times employees relating to *any* decrease in revenue or web traffic, regardless of the scale, cause, or relation to Defendants' products. *Cf. Winfield v. City of New York*, 2018 WL 2293070, at *1 (S.D.N.Y. May 18, 2018) (expressing skepticism towards requests for productions that begin "with any and all documents and communications that refer or relate to a particular subject because this type of request is contrary to Rule 34's mandate that requests must describe with reasonable particularity each item or category of items to be inspected"). Accordingly, Judge Wang correctly denied Defendants' motion for yet more discovery, as "[t]he broad scope of production sought here is simply not relevant to [their] purported fair use defense." Dkt. 344 at 5.

B. Microsoft’s Request for *All* Surveys Regarding Subscription Cancellations Is Overbroad and Irrelevant.

Microsoft also argues that Judge Wang erred by denying its motion to compel the production of *all* surveys and analyses reflecting Times subscribers’ reasons for cancelling their subscriptions. *See* Dkt. 369 at 15. It is wrong for multiple reasons.

First, The Times had already agreed to provide Defendants with the results of any survey that allows users cancelling their subscription to select, as one of the reasons for the cancellation, that “AI makes my subscription unnecessary” or that otherwise takes Generative AI into consideration. The Times is providing these results regardless of whether the user actually selected an AI-related option as the basis for their cancellation, so Microsoft’s argument that The Times is presenting only evidence favorable to its position is wrong.

Second, Microsoft’s request for surveys and other reports that do not consider or address generative AI are simply not relevant to this case. Cancellation surveys that mention generative AI are potentially relevant to show that Defendants’ products are substitutive: if a subscriber selects an option indicating that AI makes their subscription unnecessary, that indicates Defendants’ generative AI products could act as substitutes for The Times’s content. If the survey did not even present a generative AI-based reason for cancellation, however, then the responses say nothing about whether Defendants’ products are substitutive. The Times is therefore producing the only surveys relevant to the substitution inquiry—that is, those that allow a reader to point to AI products as a reason for cancellation.³

Third, Microsoft’s claim that it needs these surveys because they are purportedly “the most relevant evidence to assessing the causes of any alleged losses,” Dkt. 369 at 15, is based on a

³ Furthermore, The Times notes that it is producing documents sufficient to show trends in active Times subscribers for each month beginning in January 1, 2018, through the first quarter of 2024.

misunderstanding of The Times's claims. The Times is not seeking to attribute every subscription cancellation to Defendants. Rather, The Times has suffered subscription losses attributable to Defendants because, as a consequence of Defendants' products, users are less likely to turn to The Times's content in the first place. Reduced web traffic necessarily means fewer opportunities to convert visitors to subscribers, fewer advertising opportunities, and less affiliate link interaction. But the cancellation surveys Microsoft claims it needs have nothing to do with these lost conversion opportunities and say nothing about the harm The Times is claiming. Thus, "because Defendant[s] ha[ve] failed to demonstrate the relevance of the information sought," Judge Wang's denial of Defendants' motions to compel was not clearly erroneous or contrary to law. *See Coventry Cap. US LLC*, 333 F.R.D. at 64 (noting that "[a] finding is clearly erroneous" only "if the reviewing court is left with the definite and firm conviction that a mistake has been committed").

C. *Google Does Not Support Compelling Production of These Documents.*

Finally, Defendants hinge the remainder of their argument on a single piece of dicta in *Google LLC v. Oracle America, Inc.* In outlining the evidence the jury heard before reaching its verdict for Google, the Supreme Court acknowledged that the plaintiff's CEO had testified its losses were not "attributable to Google's development of Android." *Google*, 593 U.S. at 36. Based on this single observation, Defendants conclude that the Supreme Court endorsed the broad view that every source of a "plaintiff's alleged loss" is relevant to the fair use inquiry. However, that misreads the Supreme Court's language. The Supreme Court's observation considered the harm, or absence thereof, from the defendant, *Google's use*. *See Google*, 593 U.S. at 31. Put differently, the Supreme Court did not mandate that courts authorize discovery into other sources of the plaintiff's alleged loss (e.g., general market trends, third-party products). Consistent with *Google*,

The Times has already agreed to produce discovery regarding Defendants' products and how their unlawful use of Times content affected The Times. However, Defendants seek to go further and compel the production of materials relating to every other cause for decreases in The Times's revenue and subscriptions which, as already explained, are irrelevant to the market harm inquiry.

Thus, the Court did not clearly err in denying Defendants' traffic and subscription discovery requests, as they are irrelevant to the market substitution inquiry.

II. The Court Correctly Denied Defendants' Licensing Requests, Which Are Overbroad and Irrelevant to the Market Substitution Inquiry

Next, Defendants seek all "documents and communications related to The Times's licensing efforts, including regarding the existence (or non-existence) of a market for LLM training." *See* Dkt. 373 at 6. Again, The Times agreed to produce a substantial number of documents in response to Defendants' request, including:

1. licensing agreements related to the asserted works, regardless of whether those licenses relate to generative AI;
2. documents sufficient to show all attempts to license the asserted works for training or use in generative AI;
3. all communications related to executed and contemplated agreements to license Times works to train generative AI models;
4. all documents and communications regarding attempts to license or negotiate Times works to Defendants for use in their generative AI tools;
5. documents sufficient to show the "well-established market for The Times to provide paid access to and use of the Asserted Works;"
6. documents and communications regarding any lost market value of the asserted works as a result of the harm alleged in the complaint;
7. documents and communications regarding any injury or harm suffered as a result of the conduct alleged in the complaint; and

8. documents concerning past and projected compensation from licensing the asserted works.

However, Defendants seek more, demanding ***all*** communications regarding licensing the Asserted Works for purposes ***other than*** generative AI model training, and the existence (or lack thereof) of a “functioning market” for LLM-related licenses that are ***not*** related to generative AI training. Both requests are irrelevant to the fair use inquiry and unduly burdensome—especially given that Defendants have failed to provide a satisfactory justification for why The Times’s extensive production is inadequate to establish the market for licensing the Asserted Works or the effect of Defendants’ products on that market. *See Winfield v. City of New York*, 2018 WL 840085, at *3 (S.D.N.Y. Feb. 12, 2018) (“The party seeking to compel bears the initial burden of demonstrating relevance and proportionality.”).

Defendants’ request for internal communications regarding all licensing agreements is extraordinarily overbroad; Defendants have not and could never justify the burden of forcing The Times to turn over nearly every internal communication by its licensing team. There is simply no way that all of The Times’s internal communications or negotiations about licensing agreements, unrelated to any of the allegations in this case, could be relevant to the fair use inquiry, particularly when The Times has already agreed to produce the licenses themselves. Indeed, as Judge Wang explained in her November Order, “[t]he copyright holder’s other use or licensing of their own works to other nonparties [is] simply not as issue in the fair use determination.” Dkt. 344 at 5. And crucially, Defendants’ requests also do not pertain to the relevant market at all—viz., the licensing market for The Times’s works to train generative AI models and for generative search products—and thus cannot be relevant to their fair use defense. *See Hachette Book Grp., Inc.*, 115 F.4th at 189 (noting that, in establishing the fourth fair use factor, the defendant must prove “that the secondary use does not compete in *the relevant market*” (emphasis added)). Certainly, Defendants

have pointed to no case showing that a failure to order production of all communications related to licenses, including those in different markets, is clearly erroneous or contrary to the law.

Similarly, with respect to Defendants' request for communications regarding the existence of a market for LLM-related licenses, The Times has already agreed to produce licensing agreements for the asserted works, documents sufficient to show "the well-established market for The Times to provide paid access to and use of the Asserted Works," and all documents and communications related to both executed and contemplated agreements to license Times works to train generative AI models. This extensive discovery is more than sufficient to show the existence (or non-existence) of a market for LLM-related licenses where they are related to generative AI. Defendants' requests for additional internal communications are both irrelevant to the fair use defense and unduly burdensome. *See Winfield*, 2018 WL 840085, at *3. The market for non-generative AI-related LLM licenses is out of scope.

Thus, the Court did not clearly err in denying Defendants' request to compel documents beyond those The Times has already agreed to provide in response to Defendants' licensing requests.

III. Microsoft's Remaining Arguments Are Unavailing.

A. Microsoft's Requests are Not Relevant to Assessing the Transformativeness of Microsoft's Unauthorized Use.

Microsoft contends that the Court clearly erred in ruling that The Times's use of, and positions about, "generative AI generally" were not relevant to the transformativeness inquiry—a portion of the first factor of the fair use test. The Times has already briefed this issue in its Response to OpenAI's December 6, 2024, Rule 72(a) Objection, and incorporates all the arguments therein. *See* Dkt. 429. As set forth in that brief, the question here is whether "the copier's use" of a copyrighted work has a different purpose or character, and requires no analysis of whether

other companies who develop technology in the same field might have used the technology in a different way. *Google*, 593 U.S. at 2. Indeed, because fair use is a fact-specific inquiry, analyzing other companies' use of the technology makes no sense. *See Diabat v. Credit Suisse Grp. AG*, 2024 4252502, at *8 (S.D.N.Y. Sept. 19, 2024) (noting that the fair use "inquiry is 'fact-specific' and 'depends on all relevant circumstances'").

As Judge Wang noted in her November Order, this case "is not a referendum on the benefits of Gen AI, on Plaintiff's business practices, or about whether any of Plaintiff's employees use Gen AI at work." Dkt. 344 at 5. Instead, "[t]his case is about whether Defendant[s] trained their LLMs using Plaintiff's copyrighted material, and whether that use constitutes copyright infringement." Id. And, the fair use inquiry must focus on the defendants'—that is, OpenAI and Microsoft's—"purported use of the copyrighted works." Id. The only evidence that might be relevant to the transformative question is The Times's use of, or statements about, Microsoft's generative AI tools—and The Times has already agreed to produce that evidence. *See* Ex. 1. Therefore, "[b]ecause Defendant has not demonstrated the relevance of the [remainder of the] information sought," the Court did not clearly err in denying OpenAI's motion to compel. Dkt. 344 at 2.

B. Microsoft's Requests are Not Relevant to Show Substantial Non-Infringing Use.

Microsoft also claims that "discovery into [The Times's] uses of LLM tools is directly relevant to the defense of substantial noninfringing uses." Dkt. 369 at 24. This argument is a non-starter. This defense arises from the Supreme Court's holding that "when a product is capable of 'substantial noninfringing uses,' its manufacturer cannot be liable simply for knowing that the product can be used in a way that would constitute infringement." *EMI Christian Music Grp., Inc. v. MP3Tunes, LLC*, 844 F.3d 79, 100 (2d Cir. 2016) (quoting *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 442 (1984)). In other words, the substantial non-infringing uses

defense can negate contributory liability for infringement resulting from others' use of a defendant's allegedly infringing product. *See Sony*, 464 U.S. at 442; *see also Viacom Int'l Inc. v. Youtube Inc.*, 253 F.R.D. 256, 261–62 (S.D.N.Y. 2008) (ordering a defendant asserting the substantial non-infringing use defense to produce evidence regarding non-infringing uses of its own product).

Consistent with this ruling, The Times has already agreed to produce documents regarding its use of the **Defendants'** products for which it seeks to hold Microsoft secondarily liable. However, Microsoft cites no precedent in which a court ordered a copyright plaintiff to produce evidence regarding non-infringing uses of an entirely different product—*i.e.*, The Times's alleged use of non-Defendant generative AI tools. Nor could they. The Times's use of non-Defendant generative AI products in a non-infringing manner says nothing about whether others might use Defendants' generative AI products in a non-infringing manner. Accordingly, Judge Wang did not clearly err in ruling that Microsoft's discovery requests were irrelevant to its substantial non-infringing use defense.

This argument also fails for the independent reason that, as set forth in The Times's opposition brief to Microsoft's Motion to Dismiss, the “substantial non-infringing uses” doctrine does not apply where defendants have an “‘ongoing relationship’ with the product or its end-user.” *Arista Recs. LLC v. Usenet.com, Inc.*, 633 F. Supp. 2d 124, 156 (S.D.N.Y. 2009) (quoting *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 438 (1984)); Dkt. 76 at 10-11. Here, Microsoft offers ongoing access to products that are preloaded with copyrighted content, duplicates copyrighted content for real-time synthetic search results, and delivers copyrighted content to users. That is precisely the type of “ongoing relationship” that renders the existence of substantial non-infringing uses irrelevant. *Usenet.com*, 633 F. Supp. 2d at 155-56.

CONCLUSION

Judge Wang's November Order, as applied in her December ruling, makes clear that the focus should be on Defendants' conduct when conducting the fair use analysis. Defendants have failed to show that this finding is clearly erroneous or contrary to law. It follows that Defendants cannot show that Judge Wang's imposition of some limits on discovery into The Times, in particular where The Times has freely agreed to produce significant documentation about its licensing, revenue, subscription, and website traffic, is itself clearly erroneous or contrary to law. For the foregoing reasons, The Times respectfully requests the Court affirm Judge Wang's December 2, 2024 Orders.

Dated: January 17, 2025

/s/ Ian Crosby

Ian Crosby (*admitted pro hac vice*)
 Genevieve Vose Wallace (*admitted pro hac vice*)
 Katherine M. Peaslee (*admitted pro hac vice*)
 SUSMAN GODFREY L.L.P.
 401 Union Street, Suite 3000
 Seattle, WA 98101
 Telephone: (206) 516-3880
 Facsimile: (206) 516-3883
 icrosby@susmangodfrey.com
 gwallace@susmangodfrey.com
 kpeaslee@susmangodfrey.com

Davida Brook (*admitted pro hac vice*)
 Emily K. Cronin (*admitted pro hac vice*)
 Ellie Dupler (*admitted pro hac vice*)
 SUSMAN GODFREY L.L.P.
 1900 Ave of the Stars, Suite 1400
 Los Angeles, CA 90067
 Telephone: (310) 789-3100
 Facsimile: (310) 789-3150
 dbrook@susmangodfrey.com
 ecronin@susmangodfrey.com
 edupler@susmangodfrey.com

Elisha Barron (5036850)

Zachary B. Savage (ZS2668)
Tamar Lusztig (5125174)
Alexander Frawley (5564539)
Eudokia Spanos (5021381)
SUSMAN GODFREY L.L.P.
1301 Avenue of the Americas, 32nd Floor
New York, NY 10019
Telephone: (212) 336-8330
Facsimile: (212) 336-8340
ebarron@susmangodfrey.com
zsavage@susmangodfrey.com
tlusztig@susmangodfrey.com
afrawley@susmangodrey.com
espanos@susmangodfrey.com

Scarlett Collings (*admission pending*)
SUSMAN GODFREY L.L.P.
1000 Louisiana, Suite 5100
Houston, TX 77002
Telephone: (713) 651-9366
Facsimile (713) 654-6666
scollings@susmangodfrey.com

Steven Lieberman (SL8687)
Jennifer B. Maisel (5096995)
Kristen J. Logan (*admitted pro hac vice*)
ROTHWELL, FIGG, ERNST & MANBECK, P.C.
901 New York Avenue, N.W., Suite 900 East
Washington, DC 20001
Telephone: (202) 783-6040
Facsimile: (202) 783 6031
slieberman@rothwellfigg.com
jmaisel@rothwellfigg.com
klogan@rothwellfigg.com

Attorneys for Plaintiff
The New York Times Company